

Reliability of the Brazilian Portuguese version of the Gross Motor Function Measure in children with cerebral palsy

Kênnea M. Almeida¹, Karolina A. Albuquerque¹, Marina L. Ferreira²,
Stéphanie K. B. Aguiar², Marisa C. Mancini³

ABSTRACT | Objective: To test the intra- and interrater reliability of the Brazilian Portuguese version of the 66-item Gross Motor Function Measure (GMFM-66). **Method:** The sample included 48 children with cerebral palsy (CP), ranging from 2-17 years old, classified at levels I to IV of the Gross Motor Function Classification System (GMFCS) and four child rehabilitation examiners. A main examiner evaluated all children using the GMFM-66 and video-recorded the assessments. The other examiners watched the video recordings and scored them independently for the assessment of interrater reliability. For the intrarater reliability evaluation, the main examiner watched the video recordings one month after the evaluation and re-scored each child. We calculated reliability by using intraclass correlation coefficients (ICC) with their respective 95% confidence intervals. **Results:** Excellent test reliability was documented. The intrarater reliability of the total sample was ICC=0.99 (95% CI 0.98-0.99), and the interrater reliability was ICC=0.97 (95% CI 0.95-0.98). The reliability across GMFCS levels ranged from ICC=0.92 (95% CI 0.72-0.98) to ICC=0.99 (95% CI 0.99-0.99); the lowest value was the interrater reliability for the GMFCS IV group. Reliability in the five GMFM dimensions varied from ICC=0.95 (95% CI 0.93-0.97) to ICC=0.99 (95% CI 0.99-0.99). **Conclusion:** The Brazilian Portuguese version of the GMFM-66 showed excellent intra- and interrater reliability when used in Brazilian children with CP levels GMFCS I to IV. **Keywords:** cerebral palsy; reproducibility of results; performance evaluation; rehabilitation; physical therapy.

BULLET POINTS

- The GMFM is a very important test for the area of child motor rehabilitation.
- Every new version of a measuring instrument must be tested for reliability.
- Brazilian therapists can use the translated version of GMFM with confidence.

HOW TO CITE THIS ARTICLE

Almeida KM, Albuquerque KA, Ferreira ML, Aguiar SKB, Mancini MC. Reliability of the Brazilian Portuguese version of the gross motor function measure in children with cerebral palsy. *Braz J Phys Ther.* 2016 Jan-Feb; 20(1):73-80. <http://dx.doi.org/10.1590/bjpt-rbf.2014.0131>

● Introduction

The Gross Motor Function Measure (GMFM) is a standardized, valid, reliable, and responsive tool designed to evaluate gross motor function in children with cerebral palsy (CP)^{1,2}. This measure has been widely used in research and clinical practice in the field of child rehabilitation in different countries and has served as a reference for the development of other tests and classification systems³⁻⁸.

The GMFM aims to measure gross motor function, to help define therapeutic goals, to record changes over time, to inform caregivers of the progress made in the rehabilitation process, and to enable

the development of scientific research studies in the field^{1,2}. Following training, the measure should be used preferably by pediatric physical therapists and occupational therapists¹.

The first version of the GMFM consisted of 88 items divided into five dimensions, namely, dimension A - lying and rolling (17 items); dimension B - sitting (20 items); dimension C - crawling and kneeling (14 items); dimension D - standing (13 items); and dimension E - walking, running, and jumping (24 items)¹. The score of each item is based on a four-point scale, where a score of zero (i.e. “does

¹Departamento de Educação Integrada em Saúde (DEIS), Universidade Federal do Espírito Santo (UFES), Vitória, ES, Brazil

²Curso de Fisioterapia, Universidade Federal de Minas Gerais (UFMG), Belo Horizonte, MG, Brazil

³Departamento de Terapia Ocupacional, Universidade Federal de Minas Gerais (UFMG), Belo Horizonte, MG, Brazil

Received: Dec. 22, 2014 Revised: June 16, 2015 Accepted: June 24, 2015

not initiate”) means that the child is unable to start any activity and a score of three (i.e. “*complete*”) informs that the child completes 100% of the activity tested by the item; intermediate scores (i.e. scores of 1 and 2) describe partial performances of the item¹. The GMFM includes two types of items, dynamic and static. In the dynamic items, the examiner observes the child’s movements (e.g. item 78: *standing*, where the child must kick a ball with the right foot). In the static items, the focus is on the child maintaining the initial position for a specific period of time (e.g. item 39: *the child must maintain the weight on hands and knees for 10 seconds*). The description of the expected behavior for each score is detailed in the GMFM’s manual¹. The scores of all items are added after completing the test and are converted into a percentage performance¹. The GMFM-88 version is also used to evaluate children with syndromes and other disorders affecting motor development⁹⁻¹¹.

A second version with 66 items (GMFM-66) was developed and tested using the GMFM-88 data from 537 children, making it possible to undertake Rasch analyses and extract a 66-item version^{1,12}. The original 88 items were downsized mainly by excluding items from dimensions A (reduced from 17 to four items) and B (from 20 to 15 items)¹. Rasch analysis redistributed the items over a continuous range of relative difficulty, which enables the professionals using the test to identify items within each child’s specific functional range^{12,13}. The testing procedures and scoring are conducted as in the longer version. After completing the test, the scores of all items are plotted using the Gross Motor Ability Estimator (GMAE) software, which converts the result into a scale ranging from zero to 100 and provides a map of tested items ranked by the degree of relative difficulty¹.

The GMFM-66 requires a shorter application time (approximately 45 minutes), enables the examiner to calculate the total score, even if some items are not tested, and is the recommended version for research purposes. This new version should only be used with children at levels I to IV of the Gross Motor Function Classification System (GMFCS). Children at level V should only be evaluated using the GMFM-88 because it includes more items of lower complexity^{1,2}.

The GMFM validation sample was collected in Canada and consisted of 111 children with CP, 25 with brain injury, and 34 children under five years of age with normal motor development^{1,2}. Content validity was tested by a group of experts in pediatric developmental assessment, who selected items corresponding to the

gross motor skills of a five-year-old child without motor impairment to integrate the test content¹. Inter- and intrarater reliability values were tested in 12 children with CP and evaluated by six experienced therapists. The results, assessed using the intraclass correlation coefficient (ICC), showed excellent test consistency (ICC=0.99 for both)¹. Intra- and interrater reliability values were also assessed in each test dimension, with results ranging from 0.87 to 0.99; the lowest values were detected in the interrater reliability of dimension A¹.

GMFM-66 shows psychometric characteristics similar to those of GMFM-88^{1,12-14}. The reliability of the new version was tested in some studies and consistently presented excellent estimates. A study conducted by the authors of the instrument² on 19 children with CP, who were all evaluated by the same rater, reported an intrarater reliability of ICC=0.99. Wei et al.¹⁵ assessed the GMFM-66 intra- and interrater reliability values of two raters in a sample including 20 children with CP aged between zero and three years and also observed excellent reliability, with ICC=0.97 and ICC=0.98, respectively.

The first requirement for a good standardized test is reliability, that is, the extent to which a measure is consistent and error-free, without which reliable data cannot be collected and inferences cannot be made from the data^{16,17}. Reliability is not a fixed feature but is rather the product of interactions between instruments, raters and subjects in the evaluation context¹⁸. The main types of reliability reported in the literature are interrater reliability, which is the estimate of how consistent the test is when applied by different raters using the same scale to assess the same subjects or objects, and intrarater reliability, which informs on the consistency of scores when individuals are assessed on two or more occasions by the same rater, using the same scale¹⁶⁻¹⁸.

The authors are responsible for providing the initial test validity and reliability data. However, those characteristics are never assessed definitively, as continuous evaluations of psychometric properties are required^{17,19}. Evidence of the validity and reliability of an instrument does not guarantee that it will be used validly and reliably, particularly in a population culturally different from the population in which it was developed¹⁷. Thus, an instrument’s psychometric properties must be tested each time a scale is used in a new context or with a different group of people.

The GMFM manual and score sheet were translated into Brazilian Portuguese²⁰. To the best of our knowledge, the psychometric properties of the translated instrument have not been tested yet in the Brazilian population, though the original untranslated version is often used in research studies conducted in Brazil²¹. The assessment of GMFM reliability may show local parameters for the use of the instrument in clinical practice and research studies conducted in the country. This study aimed to assess the intra- and interrater reliability values of the Brazilian Portuguese version of the GMFM-66 in CP children with GMFCS levels ranging from I to IV.

● Method

Sample

A total of 48 children diagnosed with CP were selected from clinics and rehabilitation centers between January 2013 and July 2014. The study was approved by the Human Research Ethics Committee of Universidade Federal de Minas Gerais (UFMG), Belo Horizonte, MG, Brazil (protocol no. 476.437). The parents or guardians were invited to participate and were informed about the study procedures. Those who voluntarily agreed to their child's participation in the study signed an informed consent form.

Participants were children with CP diagnosis (GMFCS levels I to IV) confirmed by medical report and with the ability to understand and follow simple verbal commands. The sample size was calculated using the equation derived from Pearson's product-moment correlation, as reported by Streiner²², considering $\alpha=0.05$, a 95% confidence interval (CI) and an expected ICC >0.90 ²⁰, which resulted in 48 children, including 12 in each GMFCS group^{17,22}.

Four raters also participated in this study, including three physical therapists and one occupational therapist, who work in the field of pediatric rehabilitation. Only one of these raters was experienced in using the GMFM. Three raters had neither experience in using the instrument nor prior access to the English manual, not even during the study.

Procedures

The raters were trained in the use of the GMFM by two therapists with experience using the instrument. The training consisted of reading the manual translated into Portuguese and discussing test application videos with experienced therapists for 12 hours.

After training, one of the raters was selected to be the primary rater and to administer the GMFM to all of the children. The other raters watched the footage to assess interrater reliability. The primary rater had previously applied the GMFM-66 to children with normal motor development as a practical exercise to become more familiar with the score sheet and test application guidelines, as suggested by the manual²⁰.

The evaluations were performed in rehabilitation centers, locations well-known to the children, in spacious rooms at least five meters long, using benches, stairs, a stopwatch, sticks and drawing lines, and circles on the ground to apply specific items²⁰. The primary rater evaluated each child only once and filled out the test score sheet at the time of on-site assessment, consulting the manual whenever necessary. The children were evaluated while wearing comfortable clothes and no shoes and without using any assistive devices or orthotics. The items administered to each child were those whose completion was considered feasible by the rater, who allowed three attempts toward identifying the best performance²⁰.

All evaluations were recorded using a digital video camera according to a predefined standardized method. For most items applied, the camera was placed on a tripod and was positioned between the frontal and sagittal planes, according to the type of movement to be recorded, so that the child's entire body or body part to be examined was visible to the camera. Items requiring a wider camera angle, namely, items including walking, were recorded with the camera on the tripod, and its digital zoom was controlled by a research assistant. Descriptive data were collected on the same day for each child, including gender, age, social class, and GMFCS level. The entire procedure lasted 50 to 60 minutes²⁰.

To assess intrarater reliability, the primary rater watched the footage and filled out the score sheet again four weeks after the on-site assessment. This period was necessary to avoid recall bias of the on-site assessment scoring¹⁶. The scores obtained in the assessments were transferred into the GMAE, wherein the final score of the GMFM-66 assessment was generated for each child. To assess interrater reliability, the other three raters received the randomized videos recorded during the assessment of each child, including information on the administered items. The blinded raters, with no access to the sample data (GMFCS and age, among others), watched the footage of the 48 children selected, filled out the GMFM score sheets independently, and entered the data in the GMAE software.

Statistical analysis

All results assessed in the data collection were entered into the Statistical Package for the Social Sciences (SPSS) version 20.0. The values for intra- and interrater reliability regarding the total sample and each GMFCS group were calculated based on the final score generated using the GMAE. The reliability of each GMFM dimension was also calculated based on the net score. A measure of consistency (ICC type 3.1) with two-way mixed analysis²³ was used to assess intrarater reliability. A measure of absolute agreement with two-way random analysis between the scores obtained in the assessments conducted by the primary rater and the other raters (ICC type 2.1)²³ was applied to assess interrater reliability.

The ICC is an estimate ranging from zero (unreliable measure) to one (perfect reliability), assessed by the ratio between inter-group variance (including random error) and total variance. The ICC is a reliability parameter appropriate for measuring the agreement or consistency between two or more interval measures²⁴. The results of correlation coefficients are interpreted as follows: weak or no correlation from zero to 0.25; fair correlation from 0.25 to 0.50; moderate correlation from 0.50 to 0.75; and very good to excellent correlation for values higher than 0.75²⁵. However, the interpretation of reliability based on the ICC values varies according to the specificity of each study. A more stringent criterion is recommended, considering an ICC minimum value of 0.90, when a

test is used for clinical decision-making for individuals with specific health conditions^{18,24}.

Results

A total of 48 children with CP, aged two to 17 years old, primarily from families of lower socio-economic classes and predominantly with spastic-type of CP, participated in this study. Table 1 outlines the main descriptive characteristics of the sample according to the GMFCS level. The mean scores obtained in the first assessment performed by the primary rater in each GMFCS group indicated that the higher the group severity is, the lower the score obtained in the GMFM-66 will be.

To calculate reliability, 48 on-site assessments and 192 video reviews were performed. Table 2 shows the results of intra- and interrater reliability rates in the total sample, in the GMFCS groups, and in the five dimensions of the GMFM-66. The GMFM-66 scores showed excellent intra- and interrater reliability rates (ICC 0.99 and 0.97, respectively) when the total sample was analyzed. The reliability rates of results assessed in each GMFM-66 dimension were also considered excellent, all with ICC>0.95. The reliability remained excellent in the four GMFCS groups, with the lowest value detected for the intrarater reliability of GMFCS IV (ICC=0.92), which also showed the widest 95% CI (0.72 to 0.98).

Table 1. Descriptive characteristics of the total sample and per GMFCS group.

Description	GMFCS I	GMFCS II	GMFCS III	GMFCS IV	Total sample
Number of children	12	12	12	12	48
Age (mean)	7.75	8.58	8.42	10.41	8.8
Gender (n)					
Female	8	2	6	9	25
Male	4	10	6	3	23
Clinical Type (n)					
Spastic hemiplegia	8	2	0	0	10
Spastic diplegia	4	7	9	2	22
Spastic quadriplegia	0	1	1	9	11
Dyskinetic	0	1	1	1	3
Ataxic	0	0	1	0	1
Mixed	0	1	0	0	1
GMFM-66 mean score	78.53	63.75	53.47	41.21	-

GMFCS=Gross Motor Function Classification System; GMFM=Gross Motor Function Measure; n=frequency.

Table 2. Intra- and interrater reliability rates.

GMFM-66 structure	N	Intrarater reliability ICC _{3,1} (95%CI)	Interrater reliability ICC _{2,1} (95%CI)
GMFM-66 (GMAE) total score	48	0.99 (0.98-0.99)	0.97 (0.95-0.98)
Result per subscale			
Dimension A	48	0.98 (0.97-0.99)	0.95 (0.93-0.97)
Dimension B	48	0.99 (0.98-0.99)	0.96 (0.94-0.98)
Dimension C	48	0.99 (0.99-0.99)	0.99 (0.98-0.99)
Dimension D	48	0.99 (0.99-0.99)	0.99 (0.99-0.99)
Dimension E	48	0.99 (0.99-0.99)	0.99 (0.99-0.99)
GMFM-66 total score (GMAE) per GMFCS			
GMFCS I	12	0.98 (0.95-0.99)	0.93 (0.84-0.98)
GMFCS II	12	0.99 (0.97-0.99)	0.98 (0.96-0.99)
GMFCS III	12	0.99 (0.99-0.99)	0.99 (0.99-0.99)
GMFCS IV	12	0.92 (0.72-0.98)	0.94 (0.84-0.98)

N=number of subjects; GMFM=Gross Motor Function Measure; GMAE=Gross Motor Ability Estimator; GMFCS=Gross Motor Function Classification System; ICC=Intraclass Correlation Coefficients; CI=Confidence Interval.

● Discussion

The present study demonstrated that the Portuguese-translated version of the GMFM-66 has excellent intra- and interrater reliability when applied to Brazilian children with CP, GMFCS levels I to IV. This result indicated that the Brazilian version corroborates the reliability rates of the original version of the test, as assessed by Russell et al.^{1,2,20}.

The GMFM has been translated into different languages in various countries, and the psychometric properties of the translated versions have been consistently assessed^{4,15,26}. The interrater reliability of the Korean version was assessed in 39 children with CP by two experienced raters following training to use the test⁴. The Korean version also showed excellent interrater reliability in the GMFM dimensions, with the ICC values ranging from 0.98 in dimension A to 0.99 in dimension E⁴. Subsequently, the same authors assessed the reliability in 84 children with CP through video reviews conducted by 10 therapists who were trained for 30 hours. The authors identified excellent interrater reliability in all GMFM dimensions (ICC values ranging from 0.97 to 0.99) and excellent intrarater reliability (ICC values ranging from 0.99 to 1.00)²⁷.

Mahasup et al.²⁶ assessed the intra- and interrater reliability rates of the GMFM in 10 Thai children with CP. Three raters participated in the study, including one experienced in using the GMFM and two raters who read the manual and received training. On-site assessments were performed by the experienced rater

and were recorded to enable the other raters to review them and score the children independently using the footage. The study showed excellent intrarater (ICC=0.99) and interrater (ICC=0.93) reliability rates regarding the total sample score²⁶. The present study is similar to the others cited above^{2,4,26} regarding the sample characteristics of both children and raters and regarding the methods and results. Furthermore, consistency between the literature studies and the present study was also observed in terms of the previous training of raters and the use of the manual during the application of the test^{2,4,26,27}. Such methodological consistency enables the comparison of results.

The evaluation of child development using standardized instruments is complementary to purely observational clinical evaluations because such a method is structured toward minimizing subjective interpretations and ensuring the consistency of the results by assessing the psychometric properties of the tests¹⁹. The reliability of a test refers to its capacity for providing consistent results. Several factors may affect the agreement of scores, including the evaluation setting, the psychological status and age of the examinee, familiarity between examinee and rater and, especially, the knowledge, experience, and skills of the raters^{16,17,28}. It is important to conduct studies that evaluate different types of reliability, to cover as much as possible the sources of error^{16,17,19}. According to the methodology used in the present study each child was evaluated with the GMFM only

once, this administration was video recorded, and evaluations were scored again from the videotaped assessments. The results showed that examiners trained to look at children with CP on the GMFM items can consistently score them (i.e. interpret the performances). However, our study did not test the test-retest reliability, in which the examiner's ability to apply the test and get a similar performance in the scores would be checked by applying the GMFM more than once to the same individuals under similar conditions. Future reliability studies should investigate the examiner's skill in using the test as a source of variability for the Portuguese version.

The training of raters in the administration and scoring of observational instruments is a major factor in the consistent administration of these instruments and a key strategy for helping to minimize errors^{1,19,20}. Russell et al.²⁸ showed that raters improve the agreement of their scores after GMFM training workshops and that the training process has a greater effect on the ability to learn and administer the test than years of pediatric clinical experience. The sample of raters in the present study included four therapists; three of them, including the primary rater, had no prior experience in using the test and had worked in pediatric rehabilitation for less than one year. However, all raters were properly trained, according to instructions in the manual, by two experienced therapists, and the results indicate that the GMFM-66 can be used by new Brazilian therapists, provided they have received previous training. The study by Ko and Kim²⁷ assessed the interrater reliability of the GMFM Korean version between one experienced rater and one inexperienced rater and identified excellent reliability in all five GMFCS levels, which corroborates the results from the present study.

The high ICC values measured in this study may also be explained by the use of video footage to assess reliability. The videos enable the raters to watch the items as many times as necessary and to pause the videos to review the scoring guidelines in the manuals, resulting in greater reliability⁴. Video review is a rather objective form of evaluation, wherein the children's performance, rather than their ability to perform a task, is observed²⁰. This strategy may be used in both clinical practice and research studies aiming to increase the reliability of assessments using the GMFM.

This study aimed to assess the reliability of the GMFM-66, which is the version that best enables the longitudinal quantification of changes in gross motor function of children with CP¹²⁻¹⁴. The improvement of

the test as a longitudinal assessment tool was enabled by the Rasch analysis, which reorganized the items in a continuum of difficulty, showing the hierarchical structure of the instrument and providing information about the prior and emergent motor functions of each child¹². The Rasch analysis also allowed one of the requirements for using parametric statistical tests in research studies to be met by transforming the GMFM-66 total score into an interval scale¹². However, the GMFM-66 may only be administered in children with CP, preferably with GMFCS levels I to IV, as in this study sample. The reliability of the GMFM Brazilian version should also be tested encompassing all test items by applying the GMFM-88 to children with different health conditions to whom this version of the test may be applied.

The original GMFM version has been used in Brazilian research studies mainly aimed at assessing the effects of interventions on the motor functions of children with CP^{21,29,30}. The translation of the manual and evidence of reliability in the Brazilian population of children with CP will enable the use of this test to be expanded to research and clinical practice. The raters reported no difficulties in using the translated versions, and the training sufficed to apply the test. However, some translation errors were identified by experienced examiners with prior knowledge of the original version. For example, the description of score 3 of one item is defined in the test score sheet as "incomplete" instead of "fully complete", which may confuse the inexperienced rater²⁰. Another issue is the lack of translation of the GMAE software into Portuguese, despite coming with the manual, which will require therapists seeking to use the GMFM-66 to have some knowledge of the English language. These and other minor adjustments may be included in the new Portuguese edition of the manual.

● Conclusion

The GMFM-66 version, translated into Portuguese, shows excellent intra- and interrater reliability values when used in children with CP between GMFCS levels I and IV and may be used in clinical practice and in research studies on Brazilian children.

● References

1. Russell DJ, Rosenbaum PL, Avery LM, Lane M. Gross motor function measure (GMFM - 66 and GMFM - 88): user's manual. London: Mac Keith Press; 2002.

2. Russell DJ, Avery LM, Rosenbaum PL, Raina PS, Walter SD, Palisano RJ. Improved scaling of the gross motor function measure for children with cerebral palsy: evidence of reliability and validity. *Phys Ther.* 2000;80(9):873-85. PMID:10960935.
3. McCarthy ML, Silberstein CE, Atkins EA, Harryman SE, Sponseller PD, Hadley-Miller NA. Comparing reliability and validity of pediatric instruments for measuring health and well-being of children with spastic cerebral palsy. *Dev Med Child Neurol.* 2002;44(7):468-76. <http://dx.doi.org/10.1111/j.1469-8749.2002.tb00308.x>. PMID:12162384.
4. Ko J, Kim M. Inter-rater reliability of the K-GMFM-88 and the GMPM for children with cerebral palsy. *Ann Rehabil Med.* 2012;36(2):233-9. <http://dx.doi.org/10.5535/arm.2012.36.2.233>. PMID:22639748.
5. Besios T, Nikolaos A, Vassilios G, Shophia B. Comparative reliability of the PEDI, GMFM and TUG Tests for children with cerebral palsy. *J Phys Ther Sci.* 2013;25(1):73-6. <http://dx.doi.org/10.1589/jpts.25.73>.
6. Yalcinkaya EY, Caglar NS, Tugcu B, Tonbaklar A. Rehabilitation outcomes of children with cerebral palsy. *J Phys Ther Sci.* 2014;26(2):285-9. <http://dx.doi.org/10.1589/jpts.26.285>. PMID:24648650.
7. Randall KE, Bartlett DJ, McCoy SW. Measuring postural stability in young children with cerebral palsy: a comparison of 2 instruments. *Pediatr Phys Ther.* 2014;26(3):332-7. <http://dx.doi.org/10.1097/PEP.0000000000000062>. PMID:24979089.
8. Rosenbaum PL, Walter SD, Hanna SE, Palisano RJ, Russell DJ, Raina P, et al. Prognosis for gross motor function in cerebral palsy: creation of motor development curves. *JAMA.* 2002;288(11):1357-63. <http://dx.doi.org/10.1001/jama.288.11.1357>. PMID:12234229.
9. Russell D, Palisano R, Walter S, Rosenbaum P, Gemus M, Gowland C, et al. Evaluating motor function in children with Down syndrome: validity of the GMFM. *Dev Med Child Neurol.* 1998;40(10):693-701. <http://dx.doi.org/10.1111/j.1469-8749.1998.tb12330.x>. PMID:9851239.
10. Ruck-Gibis J, Plotkin H, Hanley J, Wood-Dauphinee S. Reliability of the gross motor function measure for children with osteogenesis imperfecta. *Pediatr Phys Ther.* 2001;13(1):10-7. <http://dx.doi.org/10.1097/00001577-200104000-00003>. PMID:17053645.
11. Linder-Lucht M, Othmer V, Walther M, Vry J, Michaelis U, Stein S, et al. Validation of the gross motor function measure for use in children and adolescents with traumatic brain injuries. *Pediatrics.* 2007;120(4):e880-6. <http://dx.doi.org/10.1542/peds.2006-2258>. PMID:17908744.
12. Avery LM, Russell DJ, Raina PS, Walter SD, Rosenbaum PL. Rasch analysis of the gross motor function measure: validating the assumptions of the Rasch model to create an interval-level measure. *Arch Phys Med Rehabil.* 2003;84(5):697-705. [http://dx.doi.org/10.1016/S0003-9993\(02\)04896-7](http://dx.doi.org/10.1016/S0003-9993(02)04896-7). PMID:12736885.
13. Russell DJ, Leung KM, Rosenbaum PL. Accessibility and perceived clinical utility of the GMFM-66: evaluating therapists' judgments of a computer-based scoring program. *Phys Occup Ther Pediatr.* 2003;23(2):45-58. PMID:12951787.
14. Wang HY, Yang YH. Evaluating the responsiveness of 2 versions of the gross motor function measure for children with cerebral palsy. *Arch Phys Med Rehabil.* 2006;87(1):51-6. <http://dx.doi.org/10.1016/j.apmr.2005.08.117>. PMID:16401438.
15. Wei S, Su-Juan W, Yuan-Gui L, Hong Y, Xiu-Juan X, Xiao-Mei S. Reliability and validity of the GMFM-66 in 0- to 3-year-old children with cerebral palsy. *Am J Phys Med Rehabil.* 2006;85(2):141-7. <http://dx.doi.org/10.1097/01.phm.0000197585.68302.25>. PMID:16428905.
16. Portney LG, Watkins MP. *Foundations of clinical research: applications to practice.* 3rd ed. New Jersey: Pearson Education; 2008.
17. Streiner DL, Norman GR. *Health measurement scales: a practical guide to their development and use.* 4th ed. Oxford: Oxford University Press; 2008.
18. Kottner J, Audige L, Brorson S, Donner A, Gajewski BJ, Hróbjartsson A, et al. Guidelines for Reporting Reliability and Agreement Studies (GRRAS) were proposed. *J Clin Epidemiol.* 2011;64(1):96-106. <http://dx.doi.org/10.1016/j.jclinepi.2010.03.002>. PMID:21130355.
19. Richardson KP. Use of standardized tests in pediatric practice. In: Case-Smith J, editor. *Occupational therapy for children.* St Louis: Mosby; 2005. p. 217-45.
20. Russell DJ, Rosenbaum PL, Avery LM, Lane M. *Medida da função motora grossa: (GMFM - 66 & GMFM - 88): manual do usuário.* Cyrillo ST, Galvão MCS, tradutores. São Paulo: Memnon; 2011.
21. Cury VCR, Mancini MC, Fonseca ST, Melo APP, Sampaio RF, Tirado MGA. Predicting mobility gains among children with cerebral palsy after application of botulinum toxin A. *Rev Bras Fisioter.* 2009;13(1):44-51. <http://dx.doi.org/10.1590/S1413-35552009005000008>.
22. Streiner DL. Sample-size formulae for parameter estimation. *Percept Mot Skills.* 1994;78(1):275-84. <http://dx.doi.org/10.2466/pms.1994.78.1.275>.
23. Shrout PE, Fleiss JL. Intraclass correlations: uses in assessing rater reliability. *Uses in Assessing Rater Reliability Psychological Bulletin.* 1979;86(2):420-8. PMID:18839484.
24. Vet HCW, Terwee CB, Knol DL, Bouter LM. When to use agreement versus reliability measures. *J Clin Epidemiol.* 2006;59(10):1033-9. <http://dx.doi.org/10.1016/j.jclinepi.2005.10.015>. PMID:16980142.
25. Dawson B, Trapp RG. *Basic & clinical biostatistics.* 3rd ed. New York: Lange Medical Books; 2001.
26. Mahasup N, Sritisukho P, Lekskulchai R, Keawutan P. Inter-rater and intra-rater reliability of the gross motor function measure (GMFM-66) by Thai pediatric physical therapists. *J Med Assoc Thai.* 2011;94(7 Suppl 7):S139-44. PMID:22619920.
27. Ko J, Kim M. Reliability and responsiveness of the gross motor function measure-88 in children with cerebral palsy. *Phys Ther.* 2013;93(3):393-400. <http://dx.doi.org/10.2522/ptj.20110374>. PMID:23139425.
28. Russell DJ, Rosenbaum PL, Lane M, Gowland C, Goldsmith CH, Boyce WF, et al. Training users in the gross motor function measure: methodological and practical issues. *Phys Ther.* 1994;74(7):630-36. PMID:8016195.
29. Cury VCR, Mancini MC, Melo AP, Fonseca ST, Sampaio RF, Tirado MGA. The effects of the use of orthoses on the functional mobility of children with cerebral palsy. *Rev Bras Fisioter.* 2006;10(1):67-74. <http://dx.doi.org/10.1590/S1413-35552006000100009>.

30. Christovão TCL, Pasini H, Grecco LAC, Ferreira LAB, Duarte NAC, Oliveira CS. Effect of postural insoles on static and functional balance in children with cerebral palsy: a randomized controlled study. *Braz J Phys Ther.* 2015;19(1):44-51. <http://dx.doi.org/10.1590/bjpt-rbf.2014.0072>. PMID:25651134.

Correspondence

Kênea M. Almeida

Universidade Federal do Espírito Santo (UFES)
Departamento de Educação Integrada em Saúde (DEIS)
Avenida Marechal Campos, 1468, Maruípe
CEP 29043-900, Vitória, ES, Brazil
e-mail: kenea.almeida@gmail.com